Divisional Manager
Bakeman Foods Inc.

Dear Madam/ Sir,

An accurate production forecast for Feast is critical for an efficient management of production schedules, workforce allocation, inventory management and advertisement expenditures. Due to a wide variability in historical sales, forecasting future Feast sales has been a challenge and in the recent past, 50% to 200% variations were observed the actual sales and forecast. After discussion with analysts from other departments, the undersigned set up a project to develop a quantitative model for improved sales forecast.

We studied data from January 1983 to December 1987 to identify underlying patterns and key factors influencing the sales. Over the years, sales are high at the beginning of the year and low towards the end of the year. Also, sales have observed an upward movement, over the past 60 months. The monthly expenditures on non-media promotions for consumers and dealers positively influenced the monthly sales. Also, monthly promotions have a negative influence on the sales of subsequent months. Taking all these into the account, we developed a quantitative model that can forecast sales for Feast. If the budgeted amounts for Consumer Packs (250,000) and Dealer Allowance ($100,000) for January 1988 prove correct, then there is a 95% probability that the Case Shipments for January 1988 will be within 501189 ± 75595. The predictive model we have developed does not factor in the external variables such as competitors' activities.

We are enclosing a technical appendix, if you may want to refer further details. Your feedback and queries are welcome.

Sincerely,

Tarun Shrivas
Analyst, Operations Research Department

Encl. Technical Appendix

Technical Appendix

Context:

An accurate production forecast for Feast is critical for the overall business health of Bakeman Foods Inc. We have analyzed Feast's sales and promotions data of last 60 months, from Jan '83 to Dec '87, identified key factors that influence sales of Feast and developed a quantitative model to forecast Feast sales.

Data:

In the data that we analyzed, monthly sales are represented by the variable $(\boldsymbol{Case\ Shipments})_{\boldsymbol{month}}$. The non-media promotions targeted at consumers and dealers, are represented by variables $(\boldsymbol{Consumer\ Packs})_{\boldsymbol{month}}$ and $(\boldsymbol{DealerAllowance})_{\boldsymbol{month}}$ respectively. These promotions are budgeted in advance and are planned to be utilized during specific five-week promotional periods. However, going by the historical trends, some part of these promotions have continued beyond the promotional periods. As expected, promotions during a month, push the sales in that month. However, assuming that consumers' breakfast cereal consumption rates are stable and given that breakfast cereals are non-perishable items, some part of the increased sales can be attributed to inventory build-ups by consumers. As excessive inventories may take one to two months to deplete, the sales in subsequent months may decline. In other words, promotions during previous months may have a negative influence on the sales during a given month. Accordingly, we also included previous months' promotions, i.e. one-month and two-month lagged values, in our analysis. The corresponding variables are represented by $(\boldsymbol{Consumer\ Packs})_{\boldsymbol{month-1}}$, $(\boldsymbol{DealerAllowance})_{\boldsymbol{month-1}}$, $(\boldsymbol{Consumer\ Packs})_{\boldsymbol{month-2}}$ and $(\boldsymbol{DealerAllowance})_{\boldsymbol{month-2}}$.

To identify the underlying pattern in the data, we first created a plot of $(\boldsymbol{Case\ Shipments})_{\boldsymbol{month}}$ across the last 60 months (**Figure 1**).

Figure 1: Case Shipments Time Series Plot

A long-term, upward linear trend is observed in sales, over successive months. Further, over successive years, the monthly sales follow a seasonal pattern; sales are high at the beginning of a year but slowdown by year-end. To include the effect of long-term trend and the seasonality, we included $(\boldsymbol{month})$ and $(\boldsymbol{Seasonal\ Index})_{\boldsymbol{month}}$ variables in our analysis. All the variables included in our analysis, are listed in **Table 1**.

Table 1: Variables				
#	**Variable Name**	**Variable Definition**	**Sample Size**	**Time-Frame**
1	$(Case\ Shipments)_{month}$	Sales volume during the month	48	Jan'84 – Dec'87
2	$(month)$	The month during which the sales and promotions data are observed	60	Jan'83 – Dec'87
3	$(Consumer\ Packs)_{month}$	Non-media promotions during the month targeted directly at the consumers, equivalent to a 20-cent-per-package price reduction on their purchases.	60	Jan'83 – Dec'87
4	$(Consumer\ Packs)_{month-1}$	Consumer Packs during last month	59	Feb'83 – Dec'87
5	$(Consumer\ Packs)_{month-2}$	Consumer Packs during last to last month	58	Mar'83 – Dec'87
6	$(DealerAllowance)_{month}$	Non-media promotions during the given month offered to dealers for cooperative promotional efforts, equivalent to a \$4 to \$8 per-case discount on their purchases. *(each case has 24 packs)*	60	Jan'83 – Dec'87
7	$(DealerAllowance)_{month-1}$	Dealer Allowance during last month	59	Feb'83 – Dec'87
8	$(DealerAllowance)_{month-2}$	Dealer Allowance during last to last month	58	Mar'83 – Dec'87
9	$(Seasonal\ Index)_{month}$	Seasonal effect for sales	60	Jan'83 – Dec'87

Regression Model:

We used an iterative statistical regression process, stepwise regression, to identify which combination of independent variables, provides the best prediction for sales, $(Case\ Shipments)_{month}$. At the end of the iterative process, following predictive model is identified(**Table2**):

Table 2: Regression Model

$$(\boldsymbol{Case\ Shipments})_{\boldsymbol{month}} = -93190 + 1082(\boldsymbol{month}) + 0.4076 \times (\boldsymbol{Consumer\ Packs})_{\boldsymbol{month}} - 0.2170 \times (\boldsymbol{Consumer\ Packs})_{\boldsymbol{month-1}} + 0.0683(\boldsymbol{DealerAllowance})_{\boldsymbol{month}} - 0.0179 \times (\boldsymbol{DealerAllowance})_{\boldsymbol{month-2}} + 3911 \times (\boldsymbol{Seasonal\ Index})_{\boldsymbol{month}}$$

Dependent Variable: $(Case\ Shipments)_{month}$

	Parameter Estimates	**Std. Error**	**t- value**	**p-value**
(Intercept)	-93190.00	43330.00	-2.15	0.0375*
$(month)$	1082.00	408.80	2.65	0.0114*
$(Consumer\ Packs)_{month}$	0.41	0.04	9.66	0.0000***
$(Consumer\ Packs)_{month-1}$	-2.17	0.04	-5.67	0.0000***
$(DealerAllowance)_{month}$	0.07	0.01	-5.67	0.0000***
$(DealerAllowance)_{month-2}$	-0.02	0.01	-2.95	0.0052**
$(Seasonal\ Index)_{month}$	3911.00	397.20	9.85	0.0000***
Multiple R-squared	92.85%			
Adjusted R-squared	91.81%			
F- Statistic	88.79 on 6 and 41 DF, p-value: < 0.0000			

**** Significant at 0.1%*

*** Significant at 1%*

** Significant at 5%*

The parameter estimates and standard errors for dealer and consumer promotions, and their lagged values are of limited interest as they are correlated with each other. As expected, the parameter estimate for the long-term trend, $(\boldsymbol{month})$, has a positive sign. On average, controlling for seasonality and promotions, the $(\boldsymbol{Case\ Shipments})_{month}$ is estimated to increase by 1082, every subsequent month. The parameter estimate for the trend is statistically significant at 5% level of significance. The parameter estimate for the seasonality, $(\boldsymbol{Seasonal\ Index})_{month}$ has a positive sign. On average, controlling for trend and promotions, the $(\boldsymbol{Case\ Shipments})_{month}$ is estimated to increase by 3911, for every 1 index-point increase in the seasonality. The parameter estimate for seasonality is statistically significant at 1% level of significance.

Durbin-Watson test of autocorrelation:

The predictive model assumes that the residuals, difference between predicted and actual sales, do not follow any pattern i.e. there is no autocorrelation between them. We tested this assumption with Durbin-Watson test:

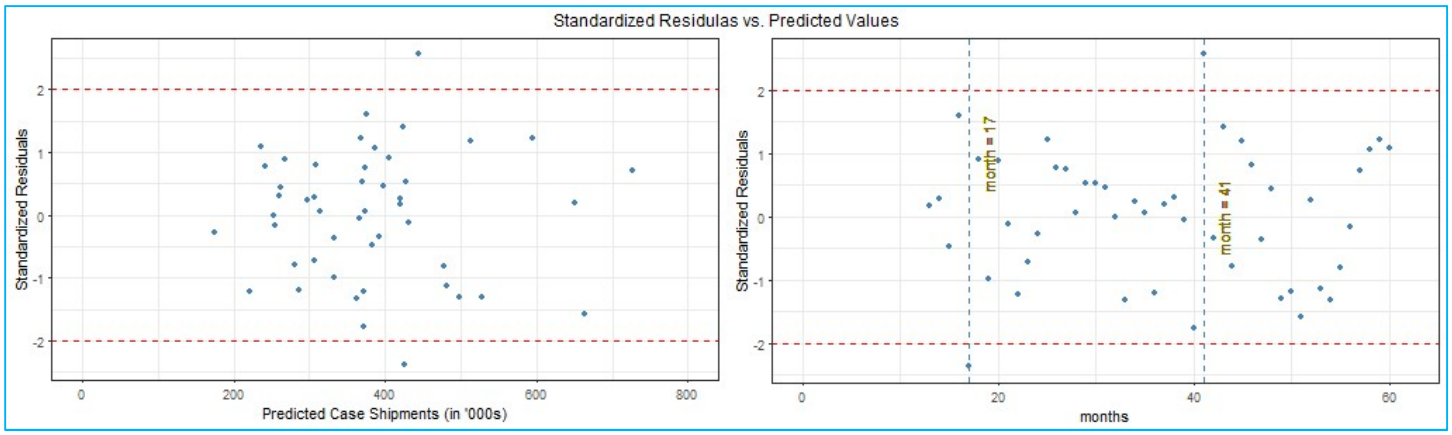
```
Durbin-Watson test

data:  model1
DW = 2.2986, p-value = 0.3701
alternative hypothesis: true autocorrelation is not 0
```

Null Hypothesis: *Autocorrelation in residuals is zero*
Alternate Hypothesis: *Autocorrelation in residuals is not zero*

A Durbin-Watson statistic of 2 corresponds to no autocorrelation amongst the residuals. With a DW statistic of 2.2986, we cannot rule out the possibility of autocorrelation. However, with a high p-value, we do not have any evidence to reject the null hypothesis, at α = 0.05. Therefore, the predictive model does present an unbiased model for the forecast.

Standardized Residuals:

The standardized residuals plot against predicted values for the Case Shipments does not have any systematic patterns and the variation in the residuals does not seem to change with the size of the predicted values.



The outliers in the residuals (month 17: May '84 & month 41: May'86) may be attributed to unusual circumstances such as plant shutdowns. The remaining residuals indicate that the predictive model has captured the patterns in the data quite well.

The goodness of Fit and Accuracy:

The Case Shipment values predicted by our model for the period between Jan 1984 (*month -13*) and Dec 1987 (month - 60) had an average accuracy of 92.85% (*Mean Average Percentage Error (MAPE) of 7.15%*). Approximately 92% of the variations *(Adjusted R-squared in Table 2)* in the Case Shipment data is explained by the predictive model. Following plot shows actual values against predicted values (**Figure 2**):

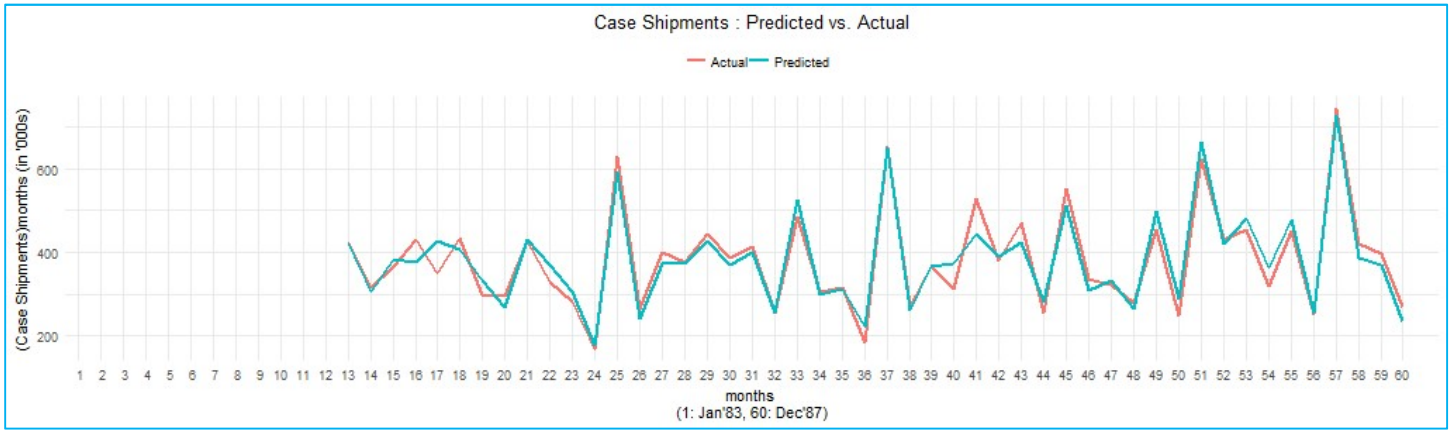


Figure 2: Case Shipments: Actual vs. Predicted

Forecast for January 1988:

Given the budgeted amounts of promotions, the Jan'1988 Case Shipments forecast are given below **(Table 3**):

Table 3: ***Forecast:*** $(Case\ Shipments)_{January\ 1988}$

Month	*Consumer Packs*	*Dealer Allowances*	*Prediction Interval*	*Point Forecast*	*Lower Limit*	*Upper Limit*	*Forecast Range*
Jan'88	250,000	$100,000	95%	501189	425594	576785	501189 ± 75595
Jan'88	250,000	$100,000	80%	501189	452433	549946	501189 ± 48757

The prediction intervals give us the uncertainty associated with the actual future values. The 95% prediction interval forecast range tells us that there is 95% probability that the Case Shipments for the month of January 1988 will be within the range 501189 ± 75595. Likewise, the forecast range for 80% prediction interval is 501189 ± 48757.

Limitations and Caveats:

Sales of Feast are also expected to be influenced by prevailing market conditions and competitors' activities. However, we have restricted our analysis to factors that we can measure or predict in advance. The forecast of Case Shipments should be interpreted, accordingly. Also, the predictive model is based on the sales and promotions data from Jan'83 to Dec'87. The model provides accurate forecasts for periods closer to this time, say next few months. On the other hand, if we are forecasting for the same month, next year, the associated variations will be more.

Word Count (excluding tables and Plots): 1158